Anderson, K.L. & C. Thiery (eds.). 2006. Information for Responsible Fisheries : Libraries as Mediators : proceedings of the 31st Annual Conference: Rome, Italy, October 10 – 14, 2005. Fort Pierce, FL: International Association of Aquatic and Marine Science Libraries and Information Centers.

GREY LITERATURE LIFE CYCLE IN RESEARCH PROJECTS: A CASE STUDY

Daniela Luzi
CNR-IRPPS

Maria Castriotta
ISPESL

Rosaria Manco and Renza Mazzucco
Link, srl.

ABSTRACT: The paper presents the results of collaboration between the Italian National Research Council (CNR) and the Italian National Institute of Safety and Health at Work (ISPESL) aiming at developing an information system - RIS-OSH (Research Information System on Occupational Safety and Health) - which gathers project descriptions and related outcomes. In the presentation of this case study, particular attention is put on the analysis of the process actually performed at ISPESL, when it funds research projects at external institutions. Workflow and information flows featuring the different steps of a project are described and some examples of the main functions of user interaction are given.

1. Introduction

The integration of information sources is one of the main concerns in the field of information and documentation [Lopatenko 2002]. Many efforts have been undertaken using technological solutions, such as the development of interoperable systems (OAI-PMH Open archive initiative - Protocol for Metadata Harvesting) and/or the adoption of standard formats such as Dublin core, XML or CERIF. In our opinion, to obtain a reliable integration it is necessary to support these efforts by improving an organizational integration. It is therefore important to reconstruct a work environment in which internal information sources are shared within an organization and in which communication among its different structures is facilitated.

The system RicercheOsh (hereafter Ris-Osh for short) was developed within the collaboration between the Italian National Research Council (CNR) and the Italian National Institute of Safety and Health at Work (ISPESL). Ris-Osh adopts this approach and integrates the types of information sources contained both in Current Research Information Systems (CRISs) [Jeffery 2002, EuroCRIS 2004] and in Institutional

Repositories (IRs) [Lynch 2003]. Moreover, we have experimented and developed the integration of Ris-Osh with other information sources produced within ISPESL.

This is possible because between CRISs and IRs there is a *relationship of continuity* due to their complementary information content. The objective of a CRIS is to reconstruct the process of research projects carried out within an institution, thus documenting the 'context' in which research outcomes are produced: the research community involved, skills and equipment exploited, budget and time employed. The description of the research project background provides an essential *preliminary* part of the IRs' information content: the management and diffusion of digital material created by the institution and its community members. As far as CRISs are concerned, this *digital material* is part of the documentation that constitutes the outcome of a research project. Moreover, digital material is mainly Grey literature (GL)[*], generally produced within research projects, where it represents the research results obtained by the project team in the form of deliverables, technical reports, guidelines, etc. For this reason we have chosen to integrate the Ris-Osh system also with the GL database, which is one of the important information sources already developed and maintained by ISPESL.

Therefore, in designing and developing Ris-Osh, the major aims were the organizational improvement of the production, collection, preservation and diffusion of institutionally generated information putting particular emphasis on the importance of the quality control of information. This aim was achieved using workflow technologies to analyze the process currently in use at ISPESL when it funds research projects at external institutions.

The paper describes some of the main issues considered in the design of Ris-Osh and shows the solutions adopted in its development through examples of interaction with the back-office of the system. The system is available at ISPESL's homepage "Research Activities" at www.ispesl.it/ricercheosh/text.

2. The design and development of the Ris-Osh system

2.1 Ris-Osh Design

The design and development of the new system is based on the description of the project life cycle using workflow methodology. Through the description of the project life cycle and of the connected information flow it was possible to identify the actors, task performed, time scheduling and exchange of information and documents of the entire process. The identification of the various actors involved in the process of supplying and

[*] According to the Luxembourg convention of 1997, recently expanded in New York 2004, Grey Literature is defined as the "Information produced on all levels of government, academics, business and industry in electronic and print formats not controlled by commercial publishing, *where publishing is not the primary activity of the producing body*".

diffusing project descriptions and related results are fundamental in developing a research information system, which can be directly updated by its information producers as soon as the information is available. This is achieved through access control and activity management of the process, which also guarantee the quality control of the information.

In analyzing the process the following organizational bottlenecks have emerged:

- ISPESL is composed of Departments, each one performing research within specific research programs, and adopting its own scientific and organizational procedures for the activation and performance of projects and for the diffusion of results.
- Project results, and in particular the final deliverables tend to remain property of one Department and there is no consolidated procedure for transmitting the documents to other departments and/or the library.
- The Library receives the deliverables resulting from the projects with a considerable delay, and the absence of a standardized format of the deliverables from external structures makes the correct bibliographical treatment of the GL documents more complex [Aceti 1999].

The development of the Ris-Osh system had, therefore, the primary aim of improving communication between departments, synchronizing and shortening the timeframes of the various activities connected to preparing and setting up projects and also improving the activities of collecting and diffusing the results. It was, however, necessary to create a new, rather flexible system that, in introducing new work organization modalities, respected the work methods and characteristics of each Department. This last requirement was satisfied by identifying the actors, for each activity, thus giving the individual departments the responsibility of assigning tasks according to their own organizational modalities.

2.2 Ris-Osh Development

Ris-Osh is composed of a back-office and a database retrievable on Internet. The back-office is based on the workflow of the project life cycle. Through an Intranet access control, the system allows the different actors to update the database with the information related to each phase of the process. The user interface facilitates the data input and provides the user with search, display and print functions for each phase of the project life cycle. It also allows the qualified user to “publish” the project information on the web, making it available to external users. The publication phase occurs after the project data have been controlled and evaluated by the staff responsible for it.

Fig. 1 – ISPESL's "Research Activities" homepage

Concerning the database schema, we have chosen to adopt the standard CERIF2000 (Common European Research Information Format) [CERIF 2000], which gives a detailed description of R&D projects and is used in CRISs. This makes the system interoperable with other CRISs. Particular attention has been paid to the choice of classifications that describe the projects. We use the Classification of Economic Activities (NACE Rev 1.1/ATECO2002) to facilitate the technological transfer and the Thesaurus OSHA to describe the aspects related to safety and health at work. Both classifications have multilingual versions and this facilitates the exchange of information also at an international level.

The system can be accessed from the ISPESL's homepage "Research Activities", which provides the user with further information on project participation, and also with a predefined format that gathers the project results from external institutions. Data from the Ris-Osh system can be retrieved using the following options: a simple textual search, a guided search, the browsing of NACE/ATECO codes or OSHA descriptors (Fig. 1).

3. Research projects life cycle

Fig. 2 describes the projects life cycle at a high level of abstraction by identifying the main phases of the process, the actors and the main activities carried out by them in each phase. Moreover, the figure shows the documents produced as a result of each activity. The actors carry out different types of activities: from the production of information and/or documents to information assessment or validation.

The events that cause the transaction from one state to the other are mainly constituted either by a positive or negative result of assessment activities performed on the quality control of information, or on its consistency. If the assessment is positive, there is a consequent permission to publish the data to users outside ISPESL. The check and assessment of the information guarantee the quality control of the data. For instance, the approval of the web publication of the deliverable represents a form of certification [Sompel 2004, Roosendaal 1997], a type of peer-review, which assesses scientific value of the deliverable, as well as the institutional choice on the opportunity to make the full-text accessible to external users.

Note that fig. 2 describes a process in which the sequence of activities is performed without any interruption, because the activities described here are always positively assessed. A more detailed description of the transactions of each phase in terms of state diagrams is given in [CNR-IRPPS 2004], [Luzi 2005].

ISPESL's funded projects are closely related to its own research activities. They are inserted into pre-defined subject areas and research programs, which are periodically established and revised, but which always reflect the institute's mission. For this reason, the ISPESL researcher, who elaborates a project proposal, supplies the potential applicant with a precise framework of the future project: a rationale and detailed objectives.

Project proposals are currently presented on a pre-printed form to the department chief, who evaluates the proposal according to the main objectives of the department, duration of the project, available funding. The researcher may be required to modify the proposal or may even have it rejected. At a set date, each department collects and sends the proposals to the ISPESL evaluation committee, which assesses them from an institutional, scientific and technical point of view. Once the proposals are approved, the administrative department integrates them into the ISPESL yearly activity plan, while the Documentation and Information Department publishes the call for proposals on Internet. At this point the phase of project assignment begins: the external applicant institutions reply to the proposal sending their project plan and description to the administrative department, which circulates them for assessment by the evaluation committee of each department. If the application is accepted, the administrative department stipulates a contract with the applicant research institution.

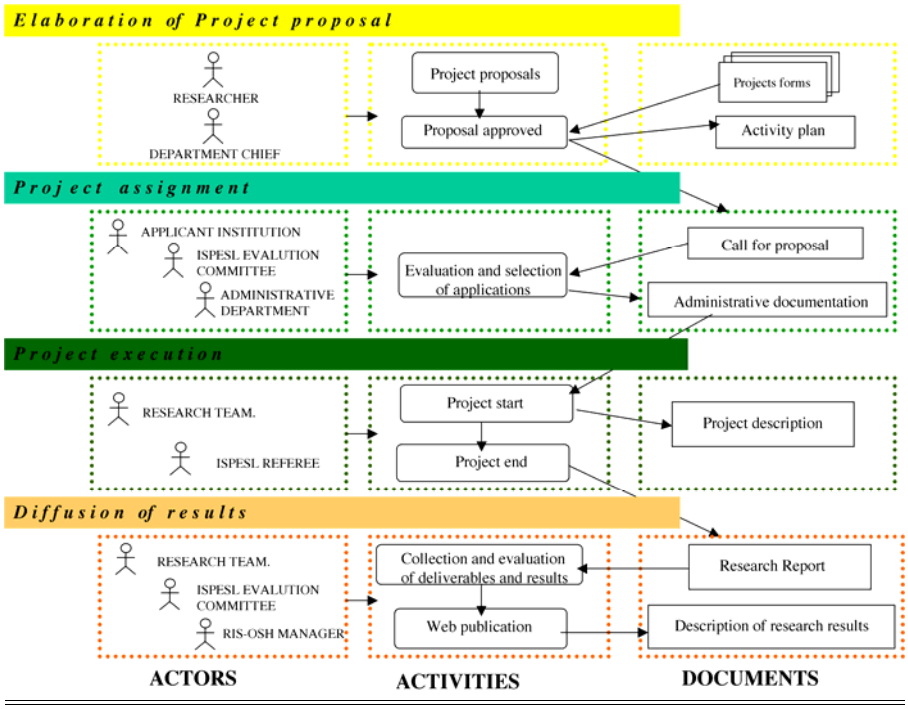

Fig. 2 - Main phases of the project life cycle

Subsequently, the project can get underway, and information about the ongoing projects can be posted on Internet to a larger target of users. At the completion of the project, the external institution provides ISPESL project referee with the results of the project described in a research report and, if required by the project, with research products (mainly databases, information guides, prototypes, safety guidelines, etc.). This documentation is collected and evaluated by ISPESL department evaluation committee. At the same time the Ris-Osh manager is notified that the project is completed through the receipt of a *description form of research results* sent electronically by the external institution. Subject to the approval of the evaluation committee, the Ris-Osh manager can now update the database with the description of research results and, if approved, with the full-text of the report. The final phase of diffusion is also marked by the integration of the Ris-Osh with the GL database. This function was developed with the aim of bettering the technological transfer represented by the improved circulation of final project reports. This integration is based on a software module, which at first controls the data and then automatically transfers the bibliographic information from Ris-Osh to the GL database.

4. Examples of system use: main functions of user interaction

The analysis of the project life cycle has shown that the optimization of the process does not result so much from its re-engineering, but rather from other features: the capability of sharing the information sources among the various actors, and a data input distributed among the identified information producers, who directly update the project information.

For this reason Ris-Osh reconstructs a workspace, both in the functionality and in the interface design, which is very similar to that used before its introduction. The similarity with the daily workflow has the advantage of facilitating user interaction with the new information tool, and at the same time, of increasing their motivation to use it.

The first layout (fig. 3) represents - in the sequence of the menus as well as graphically - the different phases of the process: the elaboration of the proposal, the editing and publication of the activity plan and of the call for proposals, the project assignment, and the collection of the research results. The only menu, which can be accessed at any phase of the process it that related to the assignment of the descriptors which classify the project: indexing can either be performed in the phase of the proposal elaboration, and/or updated by the external research institution once the project is concluded.

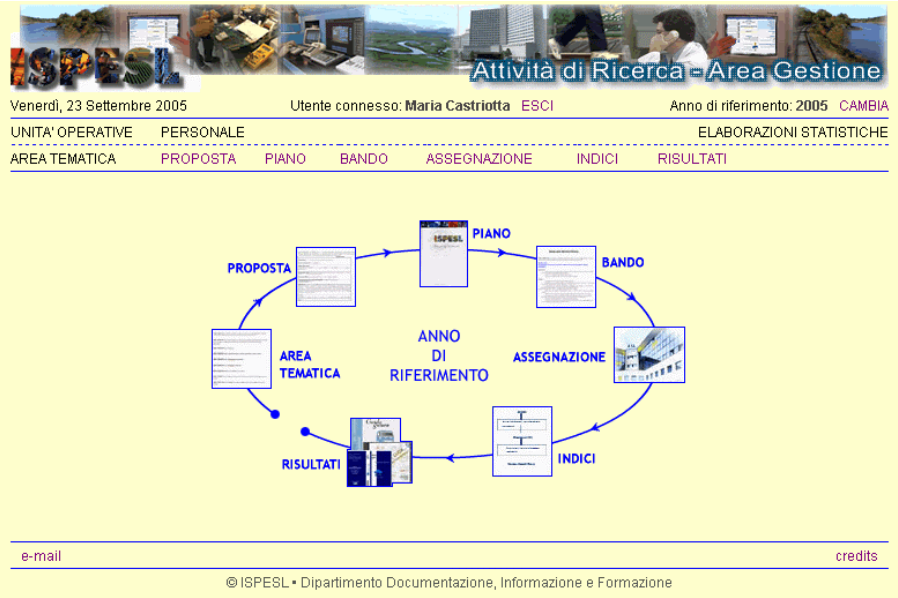

Fig. 3 – The Ris-Osh principal menus

4.1. The elaboration of the project proposal

The electronic format (Fig. 4), that the researcher fills in to make a new project proposal to the department chief, is composed of the following sections: title of the project, description of contents (objectives and rationale), project duration and costs, department and names of the proponents, notes. The researcher also connects the project proposal within the ISPESL's research program. The electronic format can be considered a personal workspace, where the researcher can elaborate his/her proposal at different times, writing, modifying or canceling parts of the information inserted. Only when all information has been filled in, the system allows the visualization and sharing of the information by the department chief.

Fig. 4 - The information describing the project proposal

4.2. The use of managerial data

The managerial data, which describe the actors of the process and their environment, are: the project proponents (researchers and affiliated department), the external institution to whom the project is assigned, the related external project manager. Figure 4 shows how to select the project proponents in the phase of the proposal elaboration, using the data taken from ISPESL personnel list. After selecting the department, the system shows the functional units and the researchers belonging to the department and allows the input of the selected names. They are visible in the right part of the module, where, if necessary, it

is possible to remove the data wrongly selected. In the central part of the module, the system shows the data in the same order of selection.

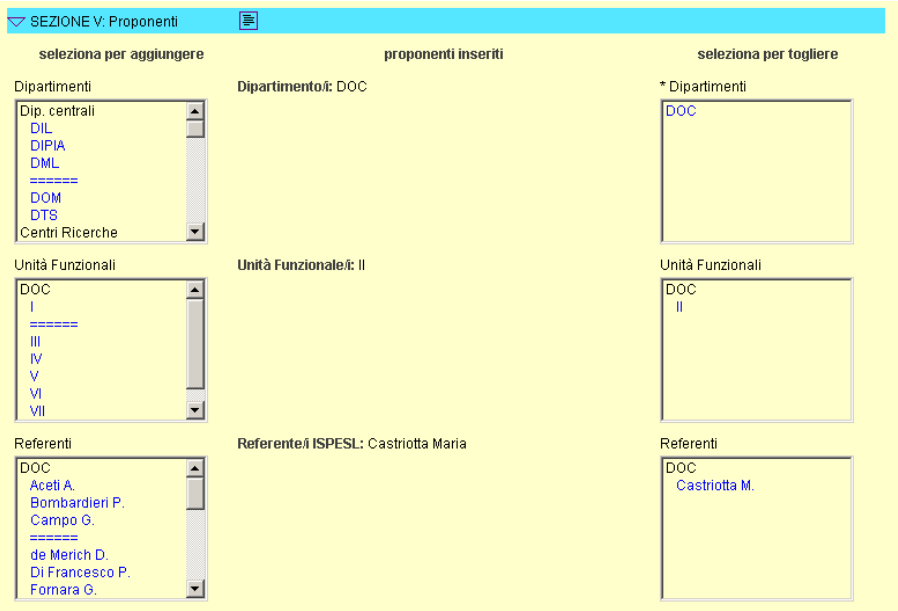

Fig. 4 – Selection and input of data related to the proposal proponents

Each role is then associated with an actor of the project: project referee, collaborator, contact person, or project manager. Figure 5 show how to associate roles to the proponents.

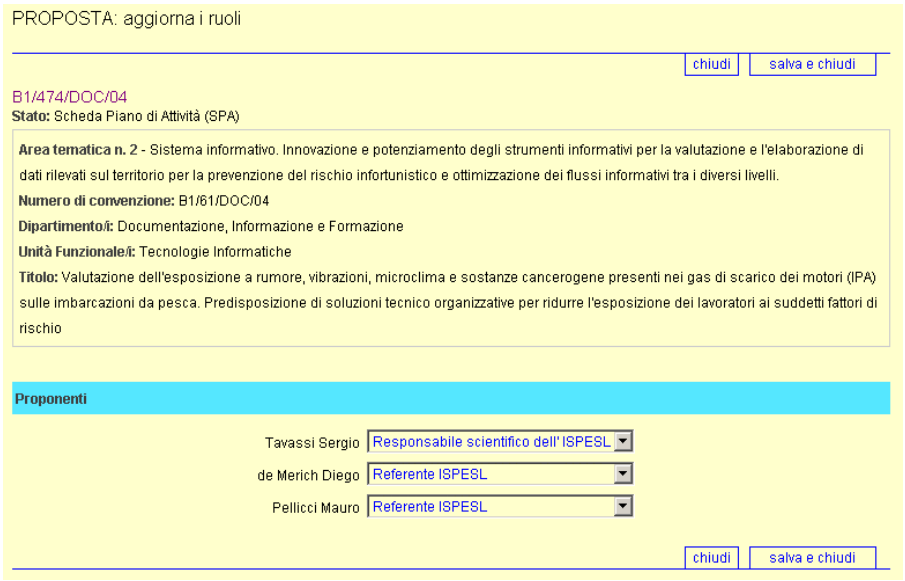

Fig. 5 – Updating of the proponents' roles

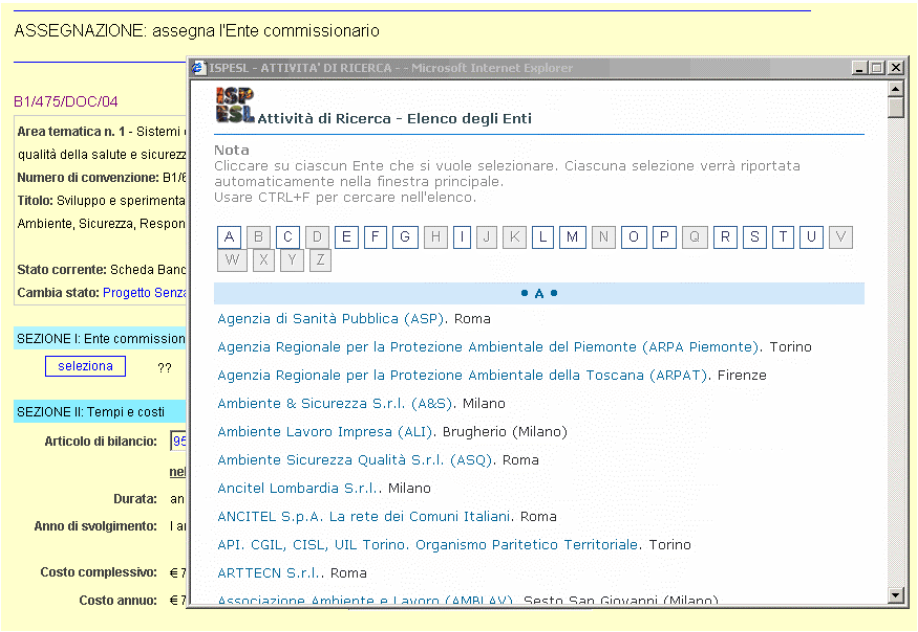

Fig. 6 – Auxiliary table for the assignment of the external institution

Managerial data related to the research team of the external institution which carries out the project are treated in the phase of project assignment (fig. 6). Clicking on the selection button, the system opens another window, which shows the alphabetic list of the external institutions that have already had a project funded by ISPESL. Selecting the institution on the window, the data is automatically inserted within the information related to the project assignment. The same approach is used to insert data about the research team.

4.3. Functions for the registration of the process status: validation and publication
The Ris-Osh system has the characteristics of making activity, which allows the transaction of states, explicit to the user. The system stores any transaction; even those that do not have a positive outcome and that therefore determine an exit point in the workflow process.

An example of this procedure is given in the assignment phase of the project life cycle. On the basis of the responses obtained to the call for proposal, three different events can occur:

1. the evaluation committee selects the external institution and assigns the project;
2. the evaluation committee considers that no response is adequate to the project objectives and does not assign the project to any external institution;
3. the call for proposals for a specific project has not received any response from any external institution and therefore the project cannot be assigned.

Figure 7 shows the simple operation which allows the user to register these events ("project without response", "project not assigned", "project assigned") and allows inputting the data related to the external institution only if the first event occurs.

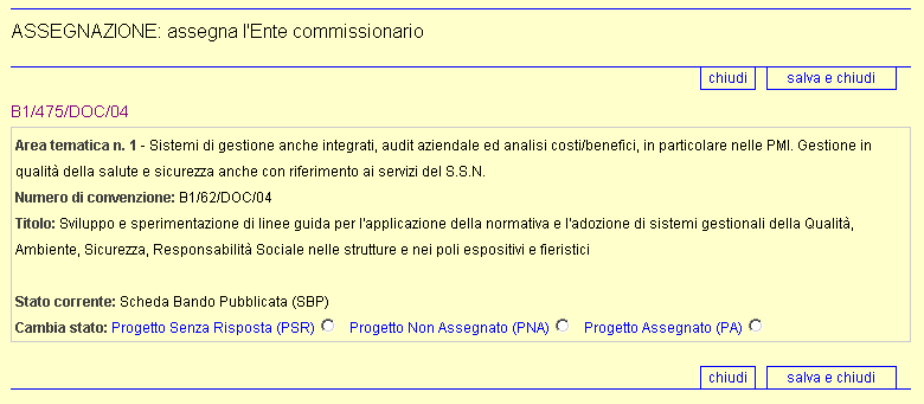

Fig. 7 - The selection of the states related to project assignment

The critical functions that determine the transaction to the next phase are mainly related to the activities of *validating* the information both from the point of view of its consistency and quality, then allowing its *publishing* on the web to users outside ISPESL.

Before publication, the system provides the user with a summary table which illustrates: a) the current state of the project life cycle; b) the number of projects which are in a specific state of the information treatment.


ASSEGNAZIONE: quadro riepilogativo di verifica per procedere alla pubblicazione

STATO CORRENTE	NUMERO SCHEDE
Progetto Assegnato (PA)	3
Scheda Bando Pubblicata (SBP)	61
Progetto Non Assegnato (PNA)	0
Progetto Senza Risposta (PSR)	0

STATO CORRENTE	NUM. MANUALE DI CONVENZIONE	NUMERO SCHEDE
PA	assegnato	3
PA	non assegnato	0

procedi alla pubblicazione

Fig. 8 Summary table giving the status of the projects before web publication

In the example showed in fig. 8 the projects contained in the call for proposal are 61; three projects have been assigned to the external institution, while there are no projects which have not received any response or which have not been assigned to any institution. In this case the publication of the data on the external institutions cannot be published because the state attributions of the majority of the projects are not finished yet. The same procedure is used to validate and publish information of the other phases of the process.

4.4. The acquisition of project results and the updating of the GL database

Figure 9 shows the acquisition module of research results, containing the editing areas to input: description of the final deliverable, of the products resulted from the project, patents, other publications (papers published in journals or conference proceedings), and notes. In the phase of results acquisition it is also possible to authorize the updating of the GL database and make the project results and, if approved, the full text of the deliverable ("elaborato finale"), available to external users on web.

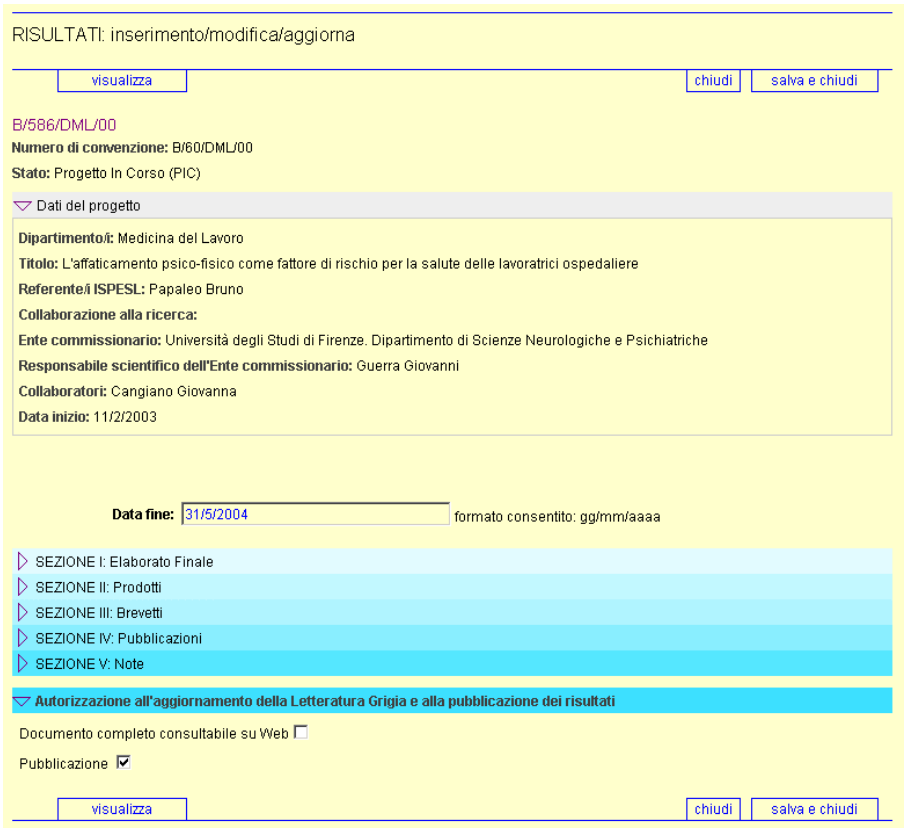

Fig. 9 - Module for the acquisition of project results

Clicking on the button for authorization of the GL database, the system sends an e-mail to the GL manager to inform that a new record has been created and reports its ID number, while the Ris-Osh manager is also notified that the bibliographic information on the deliverable has been successfully transferred.

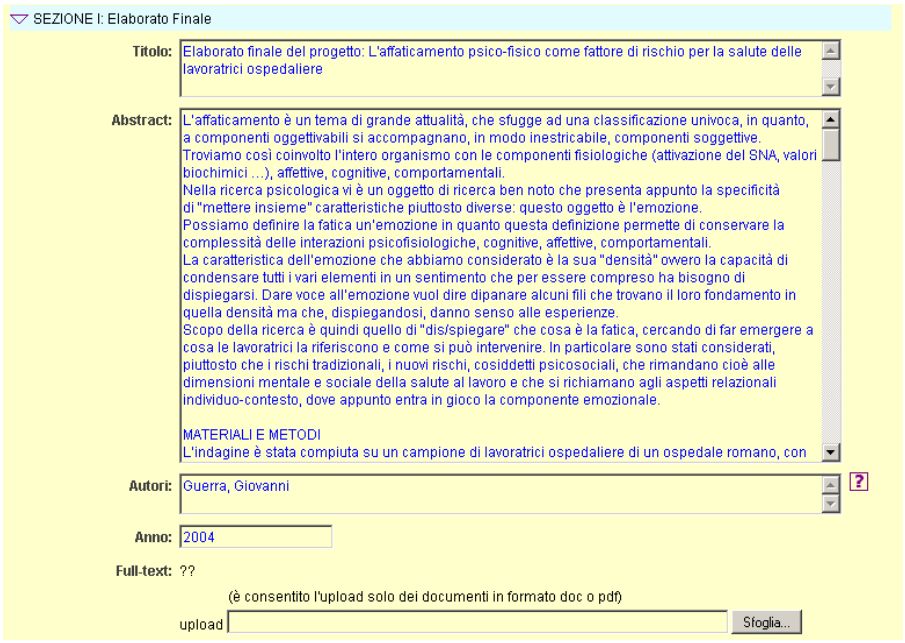

Fig. 10 –Description of the deliverable in Ris-Osh

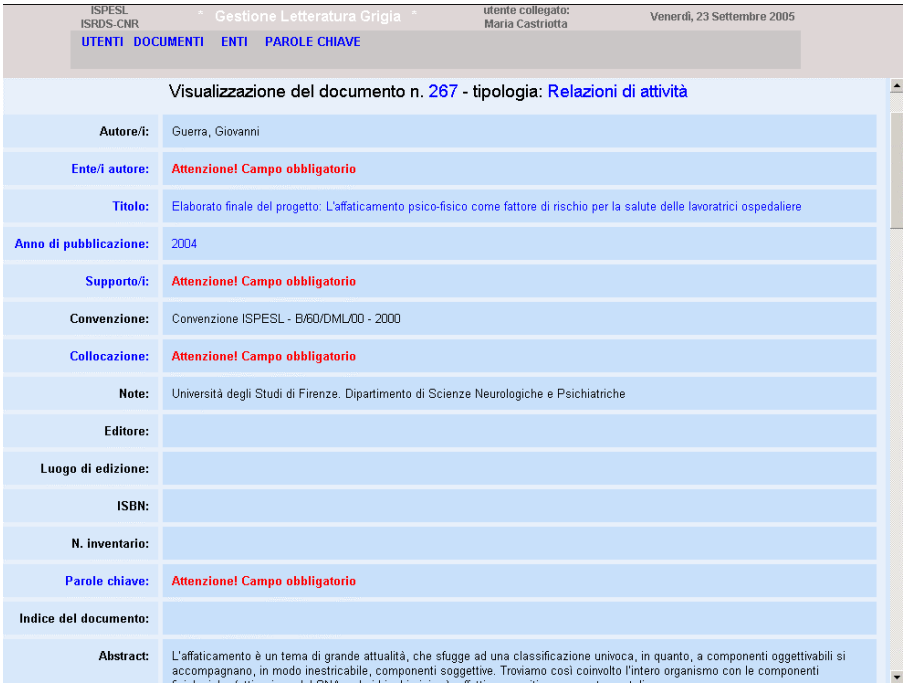

Fig. 11 – Record transferred to the GL database

Fig 10 shows the editing area for the description of the deliverable, while fig. 11 shows the record transferred to the GL database. It can be noted that the GL transferred record contains red messages that alert the GL manager to modify some information. In fact the field “corporate source” has not been filled, because this information needs a bibliographic treatment in compliance with GL cataloguing rules. That is why the Ris-Osh system transfers this information to the “note” field, compulsory data, such as location and keywords, will have to be added by the GL manager. The fields that are automatically transferred to the GL database are: the author of the deliverable, the title, the contract number, the URL to retrieve the full-text of the document, and the abstract.

4.5. Functions for the assignment of descriptors

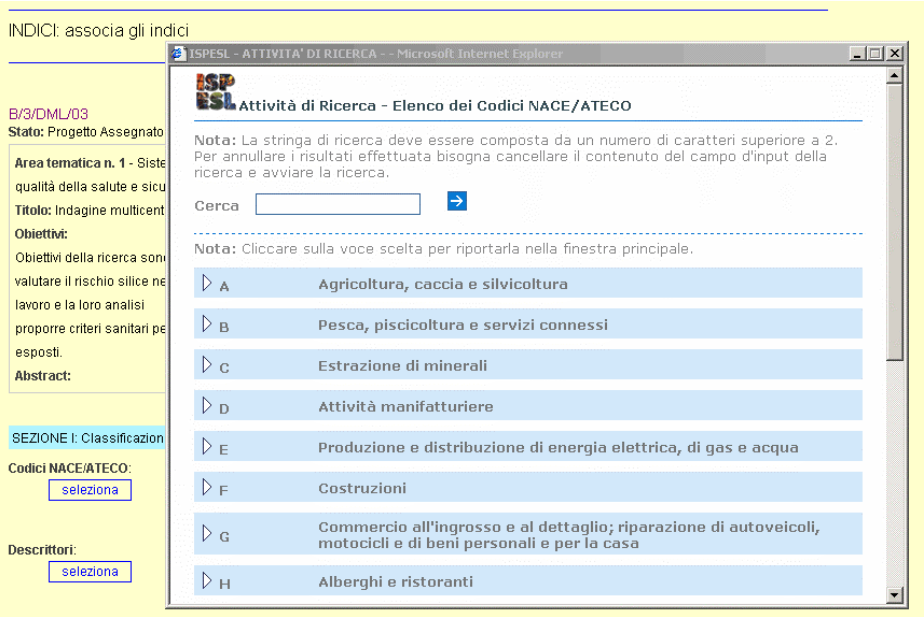

Fig. 12 – Assignment of project descriptions

As mentioned before, project descriptors can be assigned at any phase of the project life cycle, accessing the menu "Index". We use two classifications to describe the project: the Classification of Economic Activities (NACE Rev 1.1/ATECO2002), internationally used for statistical purposes by governmental R&D agencies; and the Thesaurus OSHA, which is going to be adopted by institutions in the field of safety and health at work. Figure 12 shows how the system helps the user to select one or more codes of the NACE/ATECO. The same approach is used to select descriptors from the OSHA Thesaurus. Clicking on the selection button, the system opens another window, which shows the main categories. The user can either search for the term within the classification or open the appropriate category to browse the contained codes. Selecting the code on the window, the data is automatically inserted in the module.

5. *Conclusions*

The Ris-Osh database contains now about 700 records (available at http://www.ispesl.it/ricercheosh/text) describing projects funded by ISPESL in the period 1999-2004. We are now in the phase of collecting and editing the results of the projects already concluded. At the moment the results are collected via web through an electronic format that the external institutions have to fill in when the project is concluded.

A short-term development of the Ris-Osh system concerns the optimization of the acquisition of the project results. We are planning to connect the electronic description forms with Ris-Osh allowing an automatic transfer of these data, maintaining, however, the procedures of control and evaluation of the information to be inserted.

Further improvements will regard a more extensive use of intranet and extranet access, in order to get a really integrated and shared system, where the information producers can play an active role in updating the database as soon as the information is available.

References

Aceti, A., Castriotta, M., Di Cesare, R. & Luzi, D. 1999. *A profile of GL producers in the field of safety and health at workplaces in Italy: results of a sample survey*, GL'99 Proc. On New Frontiers in Grey Literature, Washington, DC, USA, 4-5 Oct. 1999.

Common Research Information Format CERIF 2000. 1999. *CERIF 2000 Guidelines, Final Report of the CERIF Revision Working Group Co-ordinated by DG XIII-D-D*, European Commission. [Online]. Available: http://www.cordis.lu/cerif

CNR-IRPPS. 2004. Elaborato finale del progetto annuale: A*nalisi delle metodologie di trasferimento dei risultati in materia di salute e sicurezza del lavoro, mediante uso delle tecnologie di rete – Il sistema Ris-Osh***, Convenzione** B/69/DOC/00, 2004.

EuroCRIS 2004 Conference. 2004. *Putting the Sparkle in the Knowledge Society*, Antwerp, 13-15 May 2004. [Online]. Available: http://www.eurocris.org/en/taskgroups/conferences/events/cris2004/_0 .

Jeffery, K.G., Lopatenko, A. & Asserson A. 2002. *Comparative Study of Metadata for Scientific Information: The place of CERIF and Scientific Repositories*, Proc. CRIS 2002 Gaining Insight from Research Information, Aug.29-31, Kassel. [Online]. Available: http://www.uni.kassel.de/CRIS2002/programme.ghk.

Lopatenko, A., Asserson, A. & Jeffery, K.G. 2002. *CERIF, Information retrieval of Research Information in a distributed heterogeneous environment*, Proc. CRIS 2002 Gaining Insight from Research Information, Aug.29-31, Kassel. [Online]. Available: http://www.uni.kassel.de/CRIS2002/programme.ghk.

Luzi D., Castriotta M, Manco M.R., Mazzucco R. 2005. The integration of GL documents with the research information system on occupational safety and health. Proceedings of Sixth International Conference on Grey Literature, New York 6-7 December 2004. *Publishing Research Quarterly* 21 (1).

Lynch Clifford A. 2003. *Institutional Repositories: Essential Infrastructure for Scholarship in the Digital Age*, ARL Bimonthly Report 226. [Online]. Available: http://www.arl.org./newsltr/226/ir.html.

Roosendaal, Hans E., Geurts Peter A. Th. M. 1997. *Forces and functions in scientific communication: an analysis of their interplay*, CRISP 97, Cooperative Research Information Systems in Physics. [Online]. Available: http://www.physik.uni-oldenburg.de/conferences/crisp97/roosendaal.html

Van de Sompel, H., Payette, S., Erickson, J., Logoze, C. & Warner, S. 2004. Rethinking Scholarly Communication. Building the System that Scholars deserve. *D-Lib Magazine* 10 (9 Sept. 2004). [Online]. Available: http://www.dlib.org/dlib/september04/vandesompel/09vandesompel.html.